

6 Zadania

6.1 Regresja liniowa

Zaimplementuj **analityczną**(1) i **numeryczną**(2) (wykorzystując `minimize(method='Powell')`) wersje regresji liniowej. Porównaj działanie obu wersji na syntetycznym jednowymiarowym zbiorze wygenerowanym przy użyciu `sklearn.datasets.make_regression` z `noise` równym 16.

$$\vec{w} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{y} \quad (1)$$

$$L(\vec{w}) = \frac{1}{N} (\vec{y} - \mathbf{X}^T \vec{w})^T (\vec{y} - \mathbf{X}^T \vec{w}) \quad (2)$$

- $\mathbf{X} \in \mathbb{R}^{D \times N}$ próbki w zbiorze uczącym
- $\vec{y} \in \mathbb{R}^N$ etykiety zbioru uczącego
- $\vec{w} \in \mathbb{R}^D$ wagi modelu
- N ilość próbek w zbiorze / batchu

Wygeneruj trzy zbiory danych przy użyciu gry FlapPy bird:

1. Minimalny zbiór punktów zawierający tylko jeden punkt dla każdej ominiętej przeszkody.
2. Zbiór składający się z kluczowych momentów przelotu, niezbędnych do bezpiecznego pokonania trasy.
3. Kompletny zbiór składający się ze wszystkich punktów na trasie przelotu.

Parametr '`--seed`' powinien być równy numerowi indeksu. Aby zapisać aktualną pozycję postaci do pliku należy nacisnąć klawisz 'spacja'. Wykorzystując `sklearn.linear_model.LinearRegression` wykonaj regresję dla każdego ze zbiorów z wielomianem 9 i 21 rzędu (z wykorzystaniem funkcji `sklearn.preprocessing.PolynomialFeatures`) i porównaj z wynikiem własnej implementacji **analitycznej**. Dla każdego modelu oblicz błąd średniokwadratowy i narysuj dane wyjściowe oraz wynik regresji. Który model najlepiej poradził sobie z dopasowaniem do trasy?

Zaimplementuj **analityczną** wersję regresji z regularyzacją Tichonowa/L2(3) (ridge regression). Porównaj działanie własnej implementacji z `sklearn.linear_model.Ridge` na zebranych zbiorach danych z FlapPy Bird dla wielomianu 16 rzędu.

$$\vec{w} = (\mathbf{X}^T \mathbf{X} + \lambda \mathbf{I})^{-1} \mathbf{X}^T \vec{y} \quad (3)$$

- $\mathbf{I} = \text{diag}(1, 1, \dots, 1)$ macierz jednostkowa
- $\lambda \in \mathbb{R}_{>0}$ współczynnik regularyzacji

6.1.1 Kryteria oceniania

Porównanie analitycznej i numerycznej implementacji regresji na syntetycznym zbiorze $\rightarrow 3$

Porównanie `sklearn` i analitycznej implementacji regresji na zbiorze FlapPy Bird $\rightarrow 4$

Porównanie `sklearn` i własnej implementacji ridge regression na zbiorze FlapPy Bird $\rightarrow 5$